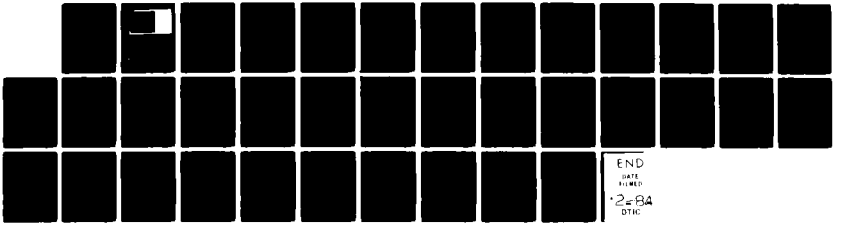
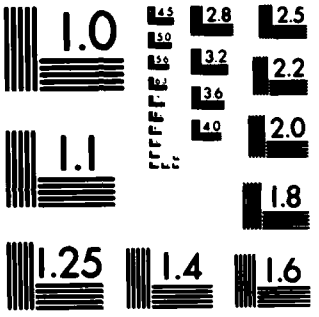
AD-A136 403 SIMPLIFIED CONFIDENCE INTERVALS FOR AND AWKWARD NONLINEAR MODEL(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER M L GOLDBERG ET AL. NOV 83 1/1

UNCLASSIFIED MRC-TSR-2601 DAAG29-80-C-0041 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

A136403

MRC Technical Summary Report #2601

SIMPLIFIED CONFIDENCE INTERVALS FOR AN AWKWARD NONLINEAR MODEL

Miriam L. Goldberg,
Peter Bloomfield
and
Margaret F. Fels

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

November 1983

(Received January 24, 1983)

DTIC
ELECTE
DEC 2 8 1983
A

DTIC FILE COPY

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

and

National Science Foundation
Washington, DC 20550

83 12 27 081

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

SIMPLIFIED CONFIDENCE INTERVALS FOR AN AWKWARD NONLINEAR MODEL

Miriam L. Goldberg*, Peter Bloomfield** and Margaret F. Fels***

Technical Summary Report #2601

November 1983

ABSTRACT

A simple method is developed for approximating confidence intervals for a linear change-point model representing residential energy consumption. The method can be applied to any nonlinear model which is conditionally linear in all but one parameter. The approximation is shown to be more accurate than the Gaussian approximation, and easier to compute than confidence intervals defined by likelihood (sum-of-squares) regions. Computations are described in some detail for the motivating model.

AMS (MOS) Subject Classification: 62J02

Key Words: Nonlinear, Regression, Confidence Intervals, Energy

Work Unit Number 4 - Statistics and Probability

---

* Mathematics Research Center, University of Wisconsin, Madison, WI 53705.

** Department of Statistics, North Carolina State University, Raleigh, NC 27650.

*** Center for Energy and Environmental Studies, Princeton University, Princeton, NJ 08544.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based upon work currently supported by the National Science Foundation under Grant No. MCS-7927062, Mod. 2, with previous funding from the Ford Foundation's State Environmental Management Program and the New Jersey Department of Energy.

SIGNIFICANCE AND EXPLANATION

A method is developed for calculating approximate confidence intervals for certain types of regression models. The confidence intervals indicate the accuracy of the parameter estimates. The method is compared with two standard procedures currently in common use. The method proposed is easier to compute than the more complicated procedure, but more accurate than the simpler one. The method is motivated and demonstrated by a simple model of residential energy consumption.

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A1	

DTIC COPY INSPECTED 3

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

SIMPLIFIED CONFIDENCE INTERVALS FOR AN AWKWARD NONLINEAR MODEL

Miriam L. Goldberg*, Peter Bloomfield** and Margaret F. Fels***

1. Introduction

Two methods are commonly used to obtain approximate confidence intervals for parameters of nonlinear models fit by the method of least squares. First is the Gaussian approximation, based on the asymptotic normality of the least squares estimate. The second is the likelihood or sum-of-squares method, based on the asymptotic chi-squared distribution of the likelihood ratio.

Asymptotically, for infinite sample sizes, the two methods are equivalent. In small samples, however, the sum-of-squares method is generally more accurate, in the sense of providing regions with coverage probabilities closer to their nominal confidence levels. The Gaussian approximation, on the other hand, based on an implicit linearization of the model function, is usually much easier to compute.

A simple procedure developed in this paper produces approximate sum-of-squares regions. These regions are in general more accurate than those based on the Gaussian approximation, but easier to compute than the actual sum-of-squares regions. An additional advantage of the method described below is that it does not require computation of derivatives of the model function.

---

* Mathematics Research Center, University of Wisconsin, Madison, WI 53705.

** Department of Statistics, North Carolina State University, Raleigh, NC 27650.

*** Center for Energy and Environmental Studies, Princeton University, Princeton, NJ 08544.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based upon work currently supported by the National Science Foundation under Grant No. MCS-7927062, Mod. 2, with previous funding from the Ford Foundation's State Environmental Management Program and the New Jersey Department of Energy.

We express the general nonlinear regression equation in matrix form, as

$$Y = \eta(\theta) + \varepsilon$$

where η is an n-dimensional vector function of the p-dimensional parameter θ, and ε is a random disturbance, also of dimension n. We assume the components ε_m to be independent, following a Normal distribution with mean 0 and variance σ^2.

Attention is restricted to model functions η which are conditionally linear in all but one parameter. That is, we assume the unknown parameter θ can be partitioned as

$$\theta = \begin{bmatrix} \underline{\lambda} \\ \tau \end{bmatrix} ,$$

with $\dim(\tau) = 1$, such that

$$\eta(\theta) = \xi(\tau)\lambda , \qquad (1.1)$$

for ξ an $n \times (p-1)$ matrix function of τ. We further confine attention to the problem of obtaining confidence intervals for individual components θ_j (i.e., for τ and for λ_j, $j = 1,2,\ldots,p-1$).

Our method is motivated and illustrated by application to a simple model of residential energy use. The energy model is described in Section 2, and the procedures used to fit the model in Section 3. In Section 4, methods of obtaining confidence intervals for the model's parameters are discussed in relation to the fitting procedure used. A simplified method of obtaining approximate confidence intervals for nonlinear models is then introduced in Section 5, and is applied to the energy model in Section 6. We conclude, in Section 7, with a discussion of the advantages of this method for general models of the form (1.1).

2. The Energy Model

For the primary heating fuel in a house, a simple model assumes that consumption on a particular day is given by the baseload α if the day's average outdoor temperature T is above a reference temperature τ, and increases in proportion to $\tau - T$ if $T < \tau$. For time period m, let Y_m denote the average daily fuel consumption, N_m the number of days, and H_m the average number of degree-days per day, to base τ:

$$H_m(\tau) = \frac{1}{N_m} \sum_{j=1}^{N_m} (\tau - T_{mj}) I(T_{mj} < \tau) \quad , \tag{2.1}$$

where I is the indicator function. The energy model is then expressed formally as

$$\begin{aligned} E(Y_m) &= \eta_m(\alpha, \beta, \tau) \\ &= \alpha + \beta H_m(\tau) \quad . \end{aligned} \tag{2.2}$$

In the terms of Equation (1.1), we have $\lambda' = [\alpha, \beta]$, where the apostrophe denotes the transpose, and $\xi(\tau) = [1, H(\tau)]$, where 1 denotes an n-vector of ones.

The daily temperature data T_{mj} are obtained in integer degrees Fahrenheit from a nearby U.S. Weather Bureau station (National Oceanic and Atmospheric Administration, monthly). The consumption data are derived from the house's monthly utility bills.

The model defined by Equation (2.2) is also used to describe aggregate data, representing a utility service region or a whole state. In this case, the observations Y_m are average daily consumption per household for each month m. In addition, the degree-day variable H_m is modified to account for the lag introduced by meters' being read on different days throughout the month. The variable H_m used for aggregate analysis in this work is defined as

$$H_m(\tau) = \frac{\sum_{j=1}^{N_m} (N_m+1-j)(\tau-T_{mj})I(T_{mj}<\tau) + \sum_{j=1}^{N_{m-1}} j(\tau-T_{m-1,j})I(T_{m-1,j}<\tau)}{\sum_{j=1}^{N_m} (N_m+1-j) + \sum_{j=1}^{N_{m-1}} j} \quad . \quad (2.3)$$

For both aggregate and individual-house analyses, a major use of the energy model is to calculate normalized annual consumption Γ. The index Γ is given by

$$\Gamma = 365(\alpha + \beta H_0(\tau)) \quad , \quad (2.4)$$

where $H_0(\tau)$ is the long-run (say, 10-year) average of daily degree-days base τ. That is, for both single and aggregate analyses, H_0 is computed from Equation (2.1), with period 0 indicating a base period of several years.

The index Γ represents the amount of fuel the house (or the average household) would consume in a year of typical temperature conditions. Comparing values of Γ obtained from two successive years of data thus offers a measure of energy conserved (or wasted) which is unbiased by the severity or mildness of a particular winter.

The theoretical basis for the model defined by Equation (2.2) is discussed by Fels and Goldberg (1982), who present analyses of aggregate residential gas consumption, based on the index Γ. Equation (2.2) and the normalized annual consumption Γ have also been used to analyze a large number of single houses heated by gas (Dutt et al, 1982), and by oil (Fels et al, 1981). Extensions of the method to electricity use in single houses are under way.

3. Fitting the Energy Model

Two basic procedures have been used to fit Equation (2.2). One is a grid search, the other an algorithm based on Newton's method. Since the model function η is linear in α and β for fixed τ, both procedures start with a search for the "best τ." Once the least squares estimate $\hat{\tau}$ is found, the estimates $\hat{\alpha}$ and $\hat{\beta}$ are found by ordinary linear regression of Y on $H(\hat{\tau})$.

The first fitting procedure, which is easier to implement, is a simple grid search taking integer ($^{\circ}F$) values of τ as grid points. The value of τ for which Y and $H(\tau)$ have the highest correlation is the best (integer) τ, and the corresponding estimates $\hat{\alpha}$ and $\hat{\beta}$ the best values of α and β for that τ. Of course, once the best integer τ is found, a more precise search can be made in a restricted neighborhood of that τ, either by taking a finer grid or by using steps from the second procedure, described below.

The grid search produces the likelihood function over the entire range of possible reference temperatures τ. This information may be of interest in a detailed analysis, but for routine applications the grid search can be too cumbersome: the variable $H(\tau)$ and its correlation with Y must be evaluated ordinarily for values of τ from $55^{\circ}F$ to $72^{\circ}F$, and in some cases for a much wider range.

The second procedure, based on Newton's method, is more efficient, but somewhat more complicated, requiring the derivatives with respect to α, β and τ of the model function η given by Equation (2.2). For this model, the derivatives are given by

$$\dot{\eta} = \frac{\partial \eta}{\partial [\alpha, \beta, \tau]} = \left[1,\ H(\tau),\ \beta \frac{dH(\tau)}{d\tau}\right] .$$

The model function η is classed as nonlinear because the derivative $\dot{\eta}$ depends on the model parameters.

In fact, the energy model is not only nonlinear, but is nondifferentiable as well: the derivative $dH/d\tau$ is discontinuous at temperature observations T_{mj}. However, except at the points $\tau = T_{mj}$ (which occur at integers) $dH/d\tau$ is well-defined. Therefore, arbitrarily choosing a right-continuous function, we define

$$\frac{dH_m(\tau)}{d\tau} \equiv F_m(\tau)$$

where, for a single house

$$F_m(\tau) = \frac{1}{N_m} \sum_{j=1}^{N_m} I(T_{mj} < \tau) \tag{3.1}$$

and for the aggregate

$$F_m(\tau) = \frac{\sum_{j=1}^{N_m} (N_m+1-j) I(T_{mj} < \tau) + \sum_{j=1}^{N_{m-1}} j \; I(T_{m-1,j} < \tau)}{\sum_{j=1}^{N_m} (N+1-j) + \sum_{j=1}^{N_{m-1}} j} \; . \tag{3.2}$$

Of course, F_m is a step function, representing the empirical distribution of the observed temperatures T_{mj}.

With the second, derivative-based procedure, the least squares estimate $\hat{\tau}$ is found by solving the normal equation

$$\{SFY - SHY \; SHF/SHH\}\big|_\tau = 0 \quad , \tag{3.3}$$

where S.. denotes a sum of squares or of cross-products, corrected for the mean. Using Newton's method, the solution is found iteratively from an initial value $\tau^{(0)}$, by setting

$$\tau^{(i+1)} = \tau^{(i)} - \left\{\frac{SFY \; SHH - SHY \; SHF}{SFY \; SHF - SHY \; SFF}\right\}\Big|_{\tau(i)} \; . \tag{3.4}$$

Since the model function η is linear between integers τ, exact convergence (within machine accuracy) occurs in one step after $\tau^{(i)}$ enters the integer interval containing the least squares estimate $\hat{\tau}$. This statement must be qualified, however, by two caveats.

We are seeking the minimum of the function $RSS(\tau)$, given by

$$RSS(\tau) = SYY - \{SHY^2/SHH\}\Big|_\tau \quad .$$

This function has a discontinuous derivative (a kink) wherever $H(\tau)$ has a discontinuous derivative - that is, at integer values of τ. Hence we have the first caveat: if $\hat{\tau}$ itself is an integer, then the minimum, $RSS(\hat{\tau})$, is at a cusp, not at a point where $dRSS/d\tau = 0$. That is, $\hat{\tau}$ will not satisfy the normal equation (3.3) in this case, so that the iteration (3.4) will never converge to $\hat{\tau}$, but will oscillate indefinitely on either side.

The second caveat is that, even when $\hat{\tau}$ does satisfy Equation (3.3), there is no guarantee that the value $\tau^{(i)}$ determined from Equation (3.4) will ever reach the interval containing $\hat{\tau}$. Instead, the iteration may oscillate over two or more intervals, excluding the right one, or may converge to a local minimum which is not the global minimum.

The possible shapes of the function $RSS(\tau)$, and the corresponding behavior of the fitting procedure, have been described in detail in Goldberg (1982). The detailed fitting algorithm, which compensates for these difficulties, is given in the Appendix. The performance of this procedure in applications to a large number of gas-heated houses is discussed in Stram et al (1982).

4. Calculating Approximate Confidence Intervals

Whatever method is used to find the least squares estimate $\hat{\tau}$, estimates of the intercept α and slope β are found by ordinary linear regression of the fuel data Y on degree-days $H(\hat{\tau})$. However, the standard errors computed from the linear regression formulae will understate the uncertainties in the estimates $\hat{\alpha}$ and $\hat{\beta}$. The understatement results from the assumption, implicit in these formulae, that the "correct" regressor $H(\hat{\tau})$ was known (or fixed) in advance, rather than being estimated from the data.

Appropriate measures of error can be obtained for the estimates of our model using methods developed for general nonlinear models. In particular, we will consider confidence intervals first based on the Gaussian approximation to the distribution of $\hat{\theta}$, and secondly based on the sum-of-squares or likelihood method. In the discussion below, these two methods are denoted by (G) for Gaussian and (S) for sum-of-squares.

The validity of applying these methods to the energy model is discussed in earlier work (Goldberg, 1982 and 1983). In the present work, we attempt to find efficient means of computing accurate confidence intervals. After summarizing the Gaussian and sum-of-squares methods, with particular reference to models of the type considered here, we will introduce two closely related approximations. From these two, we will then derive a superior, composite approximation, in Section 5. Since the sum-of-squares intervals (S) are in general much more accurate than the Gaussian (G), the former (S) are taken as the standard against which other approximations are compared.

4.1. The Gaussian Approximation

Confidence intervals based on the Gaussian approximation (G) are easily obtained for a general nonlinear model fit by the Gauss-Newton method. The variance-covariance matrix of the parameter estimate $\hat{\theta}$ is estimated by

$$\hat{V}(\hat{\theta}) = \hat{\sigma}^2(\dot{\eta}'\dot{\eta})^{-1}\Big|_{\hat{\theta}} , \tag{4.1}$$

with

$$\hat{\sigma}^2 = RSS(\hat{\theta})/(n-p) . \tag{4.2}$$

The standard error s_j of the j^{th} component of $\hat{\theta}$ is the square root of the j^{th} diagonal element of $\hat{V}(\hat{\theta})$. The confidence interval for θ_j of approximate confidence level $1 - \pi$ (for some small probability π) is therefore bounded by

$$\hat{\theta}_j \pm s_j t_{n-p}^{\pi/2} \tag{4.3}$$

where $t^{\pi/2}_{n-p}$ is the $1 - \pi/2$ quantile of the t-distribution on $n-p$ degrees of freedom.

For the energy model, when the reference temperature τ is estimated by the procedure based on Newton's method, approximate confidence intervals for α, β, and τ can be obtained from Equations (4.1) and (4.3). For the fourth parameter of interest, Γ, approximate variance and covariances are obtained using the linearization

$$\hat{\Gamma} \doteq \bar{Y} + \hat{\beta}(H_0(\tau) - \bar{H}(\tau)) + (\hat{\tau}-\tau)\beta(F_0(\tau) - \bar{F}(\tau)) \quad , \tag{4.4}$$

where $\bar{X}$ denotes the sample mean of X_m. For both single houses and aggregates, the normalized derivative F_0 is defined by Equation (3.1), with "month" 0 defined as for H_0.

With variances and covariances denoted respectively by $\hat{V}$ and $\hat{C}$, evaluating the variance-covariance matrix for the four parameters yields

$$\begin{aligned}
\hat{V}(\hat{\beta}) &= \frac{\hat{\sigma}^2}{SHH(1-r^2)} \\
\hat{V}(\hat{\tau}) &= \frac{\hat{\sigma}^2}{\beta^2 SFF(1-r^2)} \\
\hat{C}(\hat{\beta},\hat{\tau}) &= \frac{-\hat{\sigma}^2 SHF}{\beta SHH\ SFF(1-r^2)} \\
\hat{V}(\hat{\alpha}) &= \frac{\hat{\sigma}^2}{n} + \hat{V}(\hat{\beta})(\bar{H})^2 + 2\hat{C}(\hat{\beta},\hat{\tau})\bar{H}\,\bar{F} + \hat{V}(\hat{\tau})(\hat{\beta}\bar{F})^2 \\
\hat{V}(\hat{\Gamma}) &= \frac{\hat{\sigma}^2}{n} + \hat{V}(\hat{\beta})(\Delta H)^2 + 2\hat{C}(\hat{\beta},\hat{\tau})(\Delta H)\beta(\Delta F) + \hat{V}(\tau)(\beta\Delta F)^2 \\
\hat{C}(\hat{\Gamma},\hat{\tau}) &= \Delta H\ \hat{C}(\hat{\beta},\hat{\tau}) + \beta\Delta F\ \hat{V}(\hat{\tau})
\end{aligned} \tag{4.5}$$

where $r^2 = SHF^2/(SHH\ SFF)$, $\Delta H = H_0 - \bar{H}$, and $\Delta F = F_0 - \bar{F}$, all evaluated at $\hat{\tau}$. In Section 6, the Gaussian approximation represented by Equations (4.3) to (4.5) will be applied to particular data sets for the energy model. The covariances $\hat{C}$ given in Equations (4.5) are the only two that will be needed here.

Note that the formula for $\hat{V}(\hat{\Gamma})$ is the same as would be given by Equation (4.1), if the model function had been defined in terms of Γ, β, and τ as

$$\eta_m^*(\Gamma,\beta,\tau) = \Gamma + \beta(H_m(\tau)-H_0(\tau)) \quad .$$

This agreement is to be expected, since both Equation (4.4) and Equation (4.1) are based on a linearization of the model function η.

4.2. The Sum-of-Squares Method

The second basic method we will consider for obtaining approximate confidence intervals is the sum-of-squares method (S). For a general model of the form of Equation (1.1), if the "nonlinear" parameter τ is estimated by a grid search, the boundaries of a sum-of-squares region are easily computed during the estimation.

For any single parameter θ_j, the confidence region of approximate confidence level $1 - \pi$ is the projection onto the θ_j axis of the region bounded by a surface satisfying

$$\frac{RSS(\lambda,\tau) - RSS(\hat{\lambda},\hat{\tau})}{RSS(\hat{\lambda},\hat{\tau})/(n-p)} = F^{\pi}_{1,n-p} \qquad (4.6)$$

where $F^{\pi}_{1,n-p}$ is the $1 - \pi$ quantile of the F-distribution with 1 and $n - p$ degrees of freedom. Regions determined by Equation (4.6) are sometimes called "likelihood regions."

Let $\xi_\tau = \xi(\tau)$, and denote by $\hat{\lambda}_\tau$ the least-squares estimate of λ from the linear regression of Y on ξ_τ. Equation (4.6) can then be rewritten as

$$RSS(\lambda,\tau) - RSS(\hat{\lambda}_\tau,\tau) = \hat{\sigma}^2 F^{\pi}_{1,n-p} - (RSS(\hat{\lambda}_\tau,\tau) - RSS(\hat{\lambda},\hat{\tau}))$$

or

$$(\lambda-\hat{\lambda}_\tau)'(\xi_\tau',\xi_\tau)(\lambda-\hat{\lambda}_\tau) = \hat{\sigma}^2(F^{\pi}_{1,n-p} - f(\tau)) \qquad (4.7)$$

where $f(\tau) = (RSS(\hat{\lambda}_\tau,\tau) - RSS(\hat{\lambda},\hat{\tau}))/\hat{\sigma}^2$ is the F-statistic for testing the hypothesis that τ is the true value of that parameter. The left-hand side of the equation results from the fact that the model function is linear in λ for fixed τ. For any τ within the confidence region (of approximate confidence level $1 - \pi$) given by

$$f(\tau) \leq F^{\pi}_{1,n-p} \quad , \qquad (4.8)$$

the right-hand side of Equation (4.7) is positive, so that the equation has real solutions in λ. We will assume that the region in τ defined by Equation (4.8) is an interval, which we denote by $[\tau^-,\tau^+]$.

Equation (4.7) defines the cross section at τ in the λ subspace, for the surface given by Equation (4.6). Equation (4.7) is familiar from the theory of linear models, with the difference $F^{\pi}_{1,n-p} - f(\tau)$ taking the place of an F-value by itself. For any component λ_j, then, the extreme values $\lambda^+_{j\tau}$ and $\lambda^-_{j\tau}$ for which Equation (4.7) has a solution can be found using the linear theory formula

$$\lambda^{\pm}_{j\tau} = \hat{\lambda}_{j\tau} \pm \{\hat{V}_\tau(\hat{\lambda}_{j\tau})(F^{\pi}_{1,n-p} - f(\tau))\}^{1/2} \qquad (4.9)$$

where $\hat{V}_\tau(\hat{\lambda}_{j\tau})$ is the j^{th} diagonal element of

$$\hat{V}_\tau(\hat{\lambda}_\tau) = \hat{\sigma}^2(\xi_\tau'\xi_\tau)^{-1} \quad . \qquad (4.10)$$

A confidence interval $[\lambda^-_j,\lambda^+_j]$ is then obtained by maximizing (minimizing) the upper limit $\lambda^+_{j\tau}$ (lower limit $\lambda^-_{j\tau}$), over $\tau \in [\tau^-,\tau^+]$.

The matrix $\hat{V}_\tau$ of Equation (4.10) represents the variance of $\hat{\lambda}_\tau$ assuming τ to be known. Note that in Equation (4.10) the estimated sample variance $\hat{\sigma}^2$ is still given by Equation (4.2). Therefore, if a standard linear regression is used to compute $\hat{\lambda}_\tau$, the variance of $\hat{\lambda}_\tau$ as calculated by the routine must be multiplied by $RSS(\hat{\lambda},\hat{\tau})(n-p+1)/[(RSS(\hat{\lambda}_\tau,\tau)(n-p)]$ to

produce the values $\hat{V}_\tau(\hat{\lambda}_\tau)$. When $\tau = \hat{\tau}$, the multiplier reduces to a correction factor $(n-p+1)/(n-p)$ for the degrees of freedom.

4.3. Applications to the Energy Model

To motivate the shortcut confidence regions to be considered next, we apply the computational method just described to the parameters of the energy model. The known-τ variances given by Equation (4.10) become

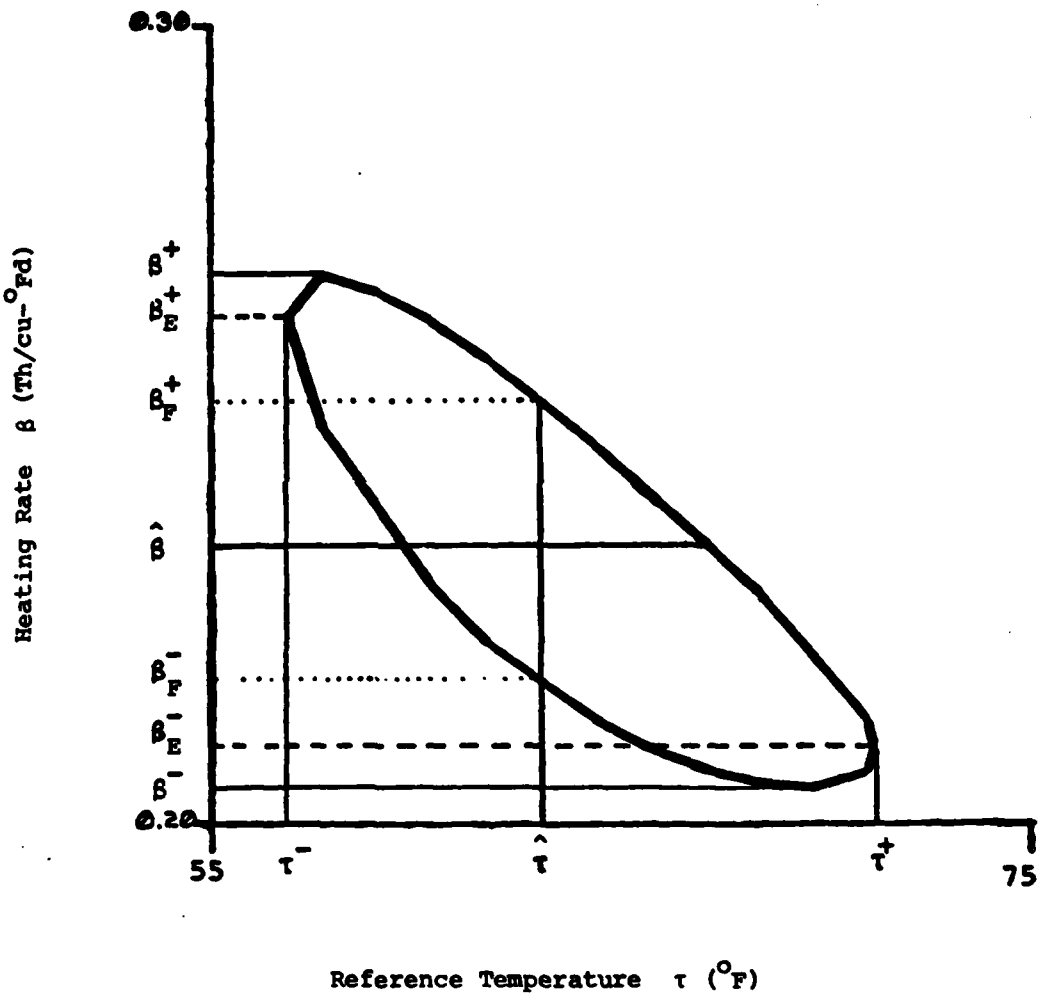
$$\begin{aligned} \hat{V}_\tau(\hat{\beta}_\tau) &= \hat{\sigma}^2/SHH \\ \hat{V}_\tau(\hat{\alpha}_\tau) &= \hat{\sigma}^2(1/n + \bar{H}^2/SHH) \\ \hat{V}_\tau(\hat{\Gamma}_\tau) &= \hat{\sigma}^2(1/n + (\bar{H}-H_0)^2/SHH) \end{aligned} \tag{4.11}$$

where H stands for $H(\tau)$. The first of Equations (4.11) was used to evaluate the contour shown in Figure 1.

This figure displays, for a particular data set fit to Equation (2.2), the projection onto the $\beta - \tau$ plane of a surface defined by Equation (4.6). The projection shown in the figure does not represent a planar cross section, with constant α, of the surface defined by Equation (4.6). Rather, for each point (β,τ) in the projection, there is some value of α such that Equation (4.6) is satisfied.

For the contour displayed, the number of observations $n = 12$, and $F^{\pi}_{1,n-p}$ was set equal to 9, corresponding to a nominal confidence level $1 - \pi$ of approximately 0.99 (for $n-p = 9$ degrees of freedom). The contour was evaluated at each integer τ satisfying Equation (4.8). In addition, the endpoints τ^- and τ^+ were found in closed form (that is, without iterative or grid search) once the integer intervals containing these points had been determined. Similarly, the β-limits β^- and β^+ were found in closed form after the extreme values for integer τ, and the corresponding integers, had been identified. The procedures used to compute these endpoints are described in Goldberg (1982, Appendix D).

Figure 1: Sum-of-Squares Contour in the β-τ Plane



Based on a fit of Equation (2.2) to data for the state aggregate, August 1969 - July 1970. The contour is defined by Equation (4.6), with $F^{\pi}_{1,9} = 9$ ($\pi = 0.01$). The abbreviations are Th for therms, cu for customer, and d for day.

The overall shape of the contour displayed in Figure 1 is roughly elliptical, but slightly bent. If the contour had been evaluated over a continuous range of τ, rather than just at integers, it would still have the same basic shape, with cusps or kinks at integer values of τ, where the function RSS has discontinuous derivatives. The segments between integers, however, would be curves, not straight lines.

4.3. Shortcut Approximations to the Sum-of-Squares Contour

Even with the simplification represented by Equation (4.9), computing the sum-of-squares intervals can be rather involved, as indicated above. To compute these intervals accurately, a very fine grid must be used, or additional steps must be taken to pinpoint the limits once they have been roughly identified. It is tempting, therefore, to look for shortcut methods to approximate a sum-of-squares region. The Gaussian approximation (G) of this region by an ellipsoid is one such shortcut. Before examining the accuracy of this approximation for the energy model (Section 6) we introduce three other shortcuts, obtained by evaluating points on the actual sum-of-squares contour.

Figure 1 shows the confidence intervals $[\beta^-,\beta^+]$ and $[\tau^-,\tau^+]$ as the projections of the sum-of-squares contour onto the corresponding axes. The figure also indicates how two of the shortcuts to be considered correspond directly to points on the contour.

First is the "fixed-τ" shortcut (F): confidence intervals for all parameters except τ are calculated as if τ had been fixed in advance. For general models of the form (1.1), the fixed-tau limits λ_{jF}^+ and λ_{jF}^- are of the form of Equation (4.3), but with $\hat{V}_\tau(\hat{\lambda})$, as given by Equation (4.10) with $\tau = \hat{\tau}$, used to estimate the variance of $\hat{\lambda}$. These limits are the points $\lambda_{j\tau}^\pm$ given by Equation (4.9) for $\tau = \hat{\tau}$, and correspond to the cross-section of

the contour at $\hat{\tau}$. For the energy model example of Figure 1, the fixed-tau interval $[\beta_F^-, \beta_F^+]$ is roughly half as wide as the sum-of-squares interval $[\beta^-, \beta^+]$.

The second shortcut displayed is the "extreme-τ" shortcut (E). For the general case, the extreme-tau limits λ_{jE}^- and λ_{jE}^+ are obtained by evaluating $\hat{\lambda}_\tau$ at the extreme values τ^- and τ^+. Implicit in this procedure is the assumption that λ_j could be determined perfectly if τ were known. In terms of the sum-of-squares contour, the points $(\tau^-, \hat{\lambda}_{\tau^-})$ and $(\tau^+, \hat{\lambda}_{\tau^+})$ are the extreme points in the τ direction. In Figure 1, the extreme-τ confidence interval $[\beta_E^-, \beta_E^+]$ understates the width of the sum-of-squares interval by only about 15%.

The fixed-tau interval (F) indicates the uncertainty the estimate $\hat{\beta}$ would have even if τ were known, while the extreme-tau interval (E) shows the uncertainty in $\hat{\beta}$ attributable to uncertainty in τ. The correct confidence interval must combine these two sources of error. As will be seen in the next section, the two simple shortcuts can be combined in a straightforward way to give a confidence interval quite close to that obtained by the complete sum-of-squares method. This composite shortcut will be shown in Section 6 to perform extremely well for all the conditionally linear parameters α, β, and Γ of the energy model.

5. The Composite Shortcut

If the energy model function η were linear in all parameters, the contour shown in Figure 1 would be a perfect ellipse. The equation for this ellipse is conveniently expressed as

$$\left(\frac{X}{X_0}\right)^2 + \left(\frac{Z}{Z_0}\right)^2 - 2\rho\left(\frac{X}{X_0}\right)\left(\frac{Z}{Z_0}\right) = 1-\rho^2 \quad , \tag{5.1}$$

where

$$X = \tau - \hat{\tau}$$

$$Z = \beta - \hat{\beta}$$

$$X_0^2 = \mathrm{Var}(\hat{\tau}) F^{*}_{1,n-p}$$

$$Z_0^2 = \mathrm{Var}(\hat{\beta}) F^{*}_{1,n-p}$$

$$\rho = \mathrm{Corr}(\hat{\beta}, \hat{\tau}) \quad .$$

For the general model of the form (1.1), the parameter β can be understood to represent any conditionally linear parameter λ_j.

An arbitrary ellipse centered at the origin can be written in the form of Equation (5.1), with X_0 and Z_0 positive. Since the equation has no real solution (X,Z) for $|X| > |X_0|$ or $|Z| > |Z_0|$, the values $\pm X_0$ and $\pm Z_0$ respectively represent the extreme values of X and Z on the ellipse.

Now assume that we know two points on the ellipse. Once is (X_0, Z_E), the single point satisfying Equation (5.1) at the extreme value $X = X_0$. The other is $(0, Z_F)$, one of the two points satisfying the equation for $X = 0$. Solving the equation for Z_E and Z_F, we find

$$Z_E = \rho Z_0 \tag{5.2}$$

and

$$\left(\frac{Z_F}{Z_0}\right)^2 = 1-\rho^2 \quad . \tag{5.3}$$

As a result, we have

$$Z_0^2 = Z_E^2 + Z_F^2 \quad . \tag{5.4}$$

Next, assume X and Z are random variables following a bivariate normal distribution with mean zero, variances $(kX_0)^2$ and $(kZ_0)^2$ respectively (for some constant k) and correlation ρ. It is a straightforward exercise to show that the conditional mean and variance of Z are given by

$$E(Z|X) = \rho \frac{X}{X_0} (Z_0) \tag{5.5}$$

$$E[(Z-E(Z|X))^2|X] = (1-\rho^2)(kZ_0)^2 \quad . \tag{5.6}$$

Comparing Equations (5.5) and (5.6) with Equations (5.2) and (5.3), we see that $(kZ_E)^2$ and $(kZ_F)^2$ respectively represent the variance of the conditional expectation of Z given X, and the expectation of the conditional variance of Z given X.

Equation (5.4) is thus a restatement of the classical decomposition of variance

$$\mathrm{Var}(Z) = \mathrm{Var}_X(E(Z|X)) + E_X(\mathrm{Var}(Z|X)) \quad . \tag{5.7}$$

Equation (5.7) may be found, for example, in Rao (1979, p. 97). This decomposition applies to general, jointly distributed random variables. The assumption of normality yields the components in proportion to ρ^2 and $1-\rho^2$.

Finally, we return to our original interpretation of X and Z as $\tau - \hat{\tau}$ and $\beta - \hat{\beta}$, respectively, for a given pair of estimates $(\hat{\tau},\hat{\beta})$. With the model function η still assumed linear, Equations (5.2) - (5.4) then become

$$\begin{aligned}(\beta^+-\hat{\beta})^2 &= \rho^2(\beta^+-\hat{\beta}) + (1-\rho^2)(\beta^+-\hat{\beta})^2 \\ &= (\beta_E^+-\hat{\beta})^2 + (\beta_F^+-\hat{\beta})^2 \quad ,\end{aligned} \tag{5.8}$$

and similarly for β^-.

The ellipse defined by Equation (5.1) now represents a confidence region for the unknown parameters β and τ, rather than a prediction region for two random variables with zero means. Nevertheless, Equation (5.8) can still be interpreted in terms of the decomposition given by Equation (5.7). A similar decomposition, given below, relates more directly to the confidence interval for β.

First, corresponding to the conditional expectation $E(Z|X)$, the least-squares estimate of β at a particular assumed value τ can be expressed as

$$\hat{\beta}_\tau = \hat{\beta} + \rho\sqrt{\frac{\mathrm{Var}\ (\hat{\beta})}{\mathrm{Var}\ (\hat{\tau})}}\,(\tau-\hat{\tau}) \quad . \tag{5.9}$$

As a result,

$$\mathrm{Var}(\hat{\beta}_\tau - \hat{\beta}) \equiv E(\hat{\beta}_\tau - \hat{\beta})^2 = \rho^2 \mathrm{Var}(\hat{\beta}) \quad . \tag{5.10}$$

If β could be determined perfectly if τ were known, as assumed by the extreme-tau method (E), then the left-hand side of Equation (5.10) would represent the variance of $\hat{\beta}$, and we would have $\rho^2 = 1$. Thus, as indicated earlier by Equation (5.2), the extreme-tau method amounts to *understating* $\mathrm{Var}(\hat{\beta})$ by a factor of ρ^2.

For the second component of the variance, corresponding to the conditional variance $\mathrm{Var}(Z|X)$, consider the variance of $\hat{\beta}_\tau$, the estimate of β when τ is known (or fixed in advance). For a linear model with uncorrelated, constant variance errors, this variance is related to the unknown-τ variance $\mathrm{Var}(\hat{\beta})$ by

$$\mathrm{Var}(\hat{\beta}_\tau) = (1-\rho^2)\mathrm{Var}(\hat{\beta}) \quad . \tag{5.11}$$

For the energy model, for example, in the case of known τ we have

$$\mathrm{Var}(\hat{\beta}_\tau) = \sigma^2/\mathrm{SHH} \quad ,$$

whereas when τ is estimated the linear approximation gives

$$\mathrm{Var}(\hat{\beta}) = \sigma^2/\mathrm{SHH}(1-\rho^2) \quad .$$

For the general case, Equations (5.10) and (5.11) together give a decomposition analogous to Equation (5.7). Assuming τ to be the true value of that parameter,

$$\mathrm{Var}(\hat{\beta}) = \mathrm{Var}(\hat{\beta}_\tau - \hat{\beta}) + \mathrm{Var}(\hat{\beta}_\tau) \quad . \tag{5.12}$$

This decomposition coincides with that given by Equation (5.7) when the estimates $\hat{\tau}$ and $\hat{\beta}$ follow a normal distribution.

Equation (5.12) can now be understood as the formal expression of an earlier statement. The extreme-tau limit β_E^+ represents the error in $\hat{\beta}$ associated with error in τ, as given by Equation (5.10). The fixed-tau

limit β_F^+ represents the error in $\hat{\beta}_\tau$ for a given τ, as in Equation (5.11). Equation (5.12) shows how the true limit β^+ combines these two sources of error, for a linear model function.

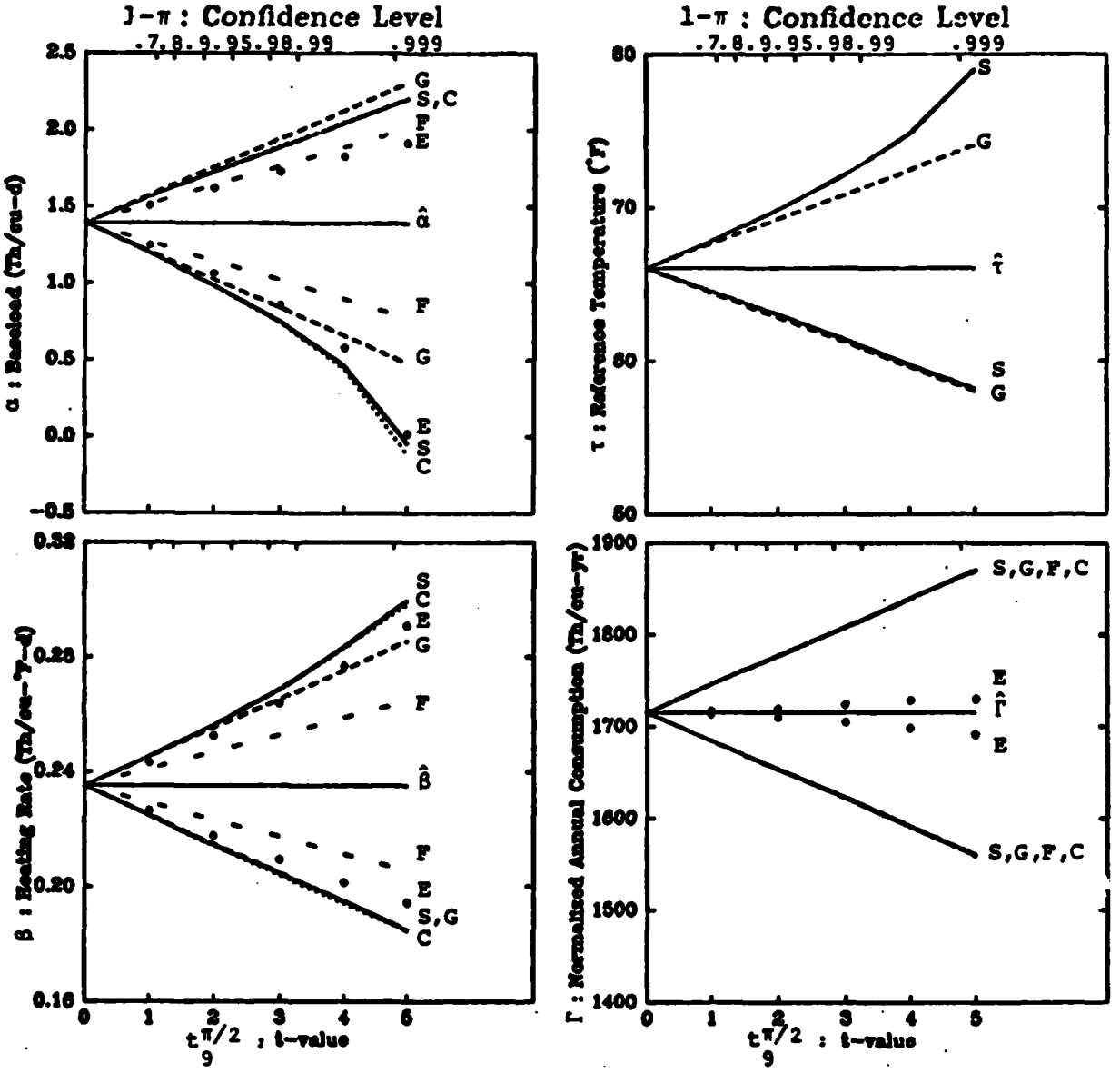
For the nonlinear model functions, the composite shortcut (C) given by Equation (5.8) offers a simple approximation to β^+. One indication that this approximation should be good for the parameters of the energy model is the fact that, for the data set of Figure 1, the shortcut limits β_E^+ and β_F^+ are quite close to those indicated by Equations (5.2) and (5.3). The correlation ρ between $\hat{\beta}$ and $\hat{\tau}$, based on Equation (4.5) for the Gaussian approximation, is -0.83 for this data set. Thus, Equations (5.2) and (5.3) imply that the extreme-tau (E) and fixed-tau (F) shortcuts should understate the sum-of-squares interval lengths (S) by 17% and 44%, respectively, as compared with the actual understatements of 15% and 45%.

As will be seen in the next section, the composite shortcut is indeed excellent for the parameters α, β, and Γ of the energy model. After examining the performance of this shortcut for our motivating model, we will turn in Section 7 to considerations for general models of the form (1.1), with an arbitrary parameter λ_j taking the place of β.

6. Application of Approximation Methods to the Energy Model

Figure 2 compares confidence intervals obtained by the sum-of-squares method (S) and by the different approximations for the data set of our earlier example. For each parameter α, β, Γ, and τ of the energy model, the different upper and lower confidence limits are plotted as functions of the nominal confidence level, or equivalently, of $t_9^{\pi/2}$. For the reference temperature τ, only the sum-of-squares (S) and Gaussian (G) methods apply. The sum-of-squares (S), extreme-tau (E), and composite (C) limits were

Figure 2. Confidence Intervals by Different Approximation Methods



The figure shows confidence intervals for the energy model parameters obtained by the Gaussian (G), Sum-of-Squares (S), Fixed-τ (F), Extreme-τ (E), and Composite (C) methods. The nominal confidence level is indicated at the top and the corresponding t-value at the bottom of each graph. See caption to Figure 1.

computed only for integer t-values. Since the Gaussian (G) and fixed-tau (F) limits are linear in the t-value, these limits are shown exactly over the range displayed.

6.1. Comparisons for Baseload, Heating Rate, and Reference Temperature

For the intercept α, slope β, and reference temperature τ, Figure 2 shows the Gaussian and sum-of-squares limits in fairly close agreement for lower confidence levels, though they diverge at higher confidence levels. For data sets like this one, then, we might be satisfied with the Gaussian approximation for any confidence level of practical interest.

However, in applications of the energy model where the standard errors of the estimates are relatively large compared to those seen here, these two sets of limits will tend to diverge earlier. In particular, for data sets with a high sample variance $\hat{\sigma}^2$ or with too few observations, the Gaussian approximation may be unreliable. Such data sets have been found occasionally in single-house analyses.

When the Gaussian approximation (G) is suspect, the composite method (C) offers a convenient improvement for the linear parameters. For both α and β, the composite limits follow the sum-of-squares limits (S) quite closely, only a slight discrepancy being seen for $t_9^{\pi/2} = 5$, corresponding to a confidence level greater than 0.999. For α as well as for β, the two components (E) and (F) of the composite approximations are of roughly equal magnitude, so that neither alone is adequate.

6.2. Approximations for Normalized Annual Consumption

A different pattern is seen for the index Γ, of major interest in energy analyses. The curves for the Gaussian (G), fixed-tau (F), and composite (C) approximations are indistinguishable from those for the sum-of-

squares limits (S) in Figure 2. The reason the four methods coincide is that the index $\hat{\Gamma}_\tau$ is quite insensitive to the nonlinear parameter τ, as indicated by the narrow bands of the extreme-tau (E) approximation. For this parameter then, it is reasonable to ignore the estimation error associated with error in $\hat{\tau}$, relying simply on the fixed-tau method (F). The composite (C) and Gaussian (G) approximations represent negligible corrections over the fixed-tau method alone. Because the accuracy of the fixed-tau method is quite useful in practice, we consider in more detail what circumstances make this method reliable.

Equations (5.2) and (5.3) indicate that more of the variance of $\hat{\Gamma}$ will be associated with the extreme-tau limits, and less with the fixed-tau limits, when the covariance ρ between $\hat{\Gamma}$ and $\hat{\tau}$ is large. The magnitude of this covariance ($\hat{C}(\hat{\Gamma},\hat{\tau})$ in Equation (4.5)) tends to be larger for larger values of ΔH and ΔF. The general insensitivity of $\hat{\Gamma}_\tau$ to τ stems from the fact that normal degree-days $H_0(\tau)$ and its derivative $F_0(\tau)$ are usually close to the corresponding sample means $\bar{H}(\tau)$ and $\bar{F}(\tau)$, so that the terms ΔH and ΔF are small.

However, the insensitivity of $\hat{\Gamma}_\tau$ to τ may not hold if the observations Y_m are not taken over an integral number of years, or are taken at unevenly spaced times. In such cases, the differences ΔH and ΔF between the base-period values and the sample means are likely to be large. Such irregularly occurring data are common, for example, for fuel oil deliveries, which are more frequent during cold months. As a result, for some oil-heated houses the correlation ρ between $\hat{\Gamma}$ and $\hat{\tau}$ is occasionally too high for the fixed-tau (F) approximation to be trusted.

For the example data set of Figure 2, $\Delta H = 1.3^{\circ}F$ and $\Delta F = 0.012$, while the correlation ρ between $\hat{\Gamma}$ and $\hat{\tau}$ is 0.07. The situation is quite different for a particular oil-heated house in the same region. For

that house, all but one of the nine observations represented winter consumption; the corresponding values are $\Delta H = 7.6^{\circ}F$, $\Delta F = 0.426$ and $\rho = 0.6$. In this case, for $\pi = 0.05$ the fixed-tau limits (F) understate the sum-of-squares limits (S) by 20% (as compared with a negligible amount for the data set in Figure 2). By contrast, the Gaussian interval of [434,938] (gallons per year) is quite close to the sum-of-squares interval [412,927]. The composite interval is even closer: [423,923].

Thus, the same relationship that was found for the parameters α and β of the ordinary aggregate gas data set displayed in Figure 2 holds also for the index Γ even in a case where the fixed-tau approximation (F) is unreliable. That is, the Gaussian approximation (G) seems to be adequate for confidence levels of practical interest, but the composite shortcut (C) is still more accurate. In the next section we will discuss why the composite should be superior in general for models of the form (1.1).

6.3. Implications for Other Energy Data Sets

A total of 75 aggregate gas data sets were fit by Equation (2.2), with the degree-day variable defined by Equation (2.3). Although they were not studied in the same detail as the data set used for Figures 1 and 2, our analysis indicates that the rest of these data sets should behave similarly. For one thing, the (Gaussian approximation) correlations among the parameter estimates $\hat{\alpha}$, $\hat{\beta}$, $\hat{\tau}$, and $\hat{\Gamma}$ were similar for all the data sets studied. Secondly, as discussed elsewhere (Goldberg, 1983), the example data set seems to be typical in terms of the severity of nonlinearities which might distort the sum-of-squares contour from an ellipse. From the previous section, we know that the success of the extreme-tau (E) and fixed-tau (F) shortcuts depends on the magnitude of the correlation ρ, while the success of the composite shortcut (C) depends on how nearly elliptical the sum-of-squares contour is.

7. Advantages of the Composite Shortcut

In applying the approximation represented by Equation (5.3) to a general model of the form (1.1), two considerations are important: ease of computation, and accuracy. As noted above, we here measure accuracy in terms of how closely the composite limits (C) approximate the sum-of-squares limits (S).

Computation of the composite limits (C) is considerably easier than finding the sum-of-squares limits (S). Once the least squares estimate $\hat{\theta}$ has been found, a nonlinear equation must be solved only to obtain the limits τ^- and τ^+ for the nonlinear parameter τ. For the linear parameters λ_j, the extreme-tau limits (E) are then easily obtained as the coefficients from a linear regression at τ^+ and τ^-, and the fixed-tau limits (F) as confidence interval bounds for a linear regression at $\hat{\tau}$.

If the model has been fit using the derivatives $\dot{\eta}$, as with the commonly used procedures based on Newton's method, the composite shortcut (C) will require more computational effort than the Gaussian approximation (G). However, for some types of models, particularly non-differentiable ones, derivative-based fitting procedures may be impractical. The complexity such methods can impose was indicated in Section 3. Even for continuously differentiable model functions η, it may be easier, and more informative, for an analyst to run a series of linear regressions, varying the parameter τ, than to implement a nonlinear regression routine correctly for his model. Provided a wide enough range of values of τ was used initially, the composite shortcut (C) can be computed directly from the output of such a series of linear regressions.

Whatever the fitting procedure used, both the Gaussian (G) and fixed-tau (F) methods offer the computational advantage of giving limits for any confidence interval of interest by multiplying a single standard error

estimate by the appropriate t-value. Further, the latter method (F) does not require the derivative of the model function η with respect to the nonlinear parameter τ. However, the linear, t-based confidence intervals (F) and (G) cannot reflect any asymmetry in the likelihood (S) region. As can be seen in Figure 2 of the previous section, this asymmetry is captured by the extreme-tau method (E), and is appropriately incorporated in the composite (C).

For the energy model, the composite approximation (C) was seen in the previous section to be quite accurate, and superior to the Gaussian approximation (G). The difficulty with the Gaussian approximation in general is that any nonlinearity in the model function will make the Gaussian ellipse different from the sum-of-squares contour, even though many types of nonlinearity will leave the sum-of-squares contour itself essentially elliptical.

Hamilton, et al (1981) presented, as an improvement over the Gaussian, an explicit elliptical approximation to the sum-of-squares regions, based on a second-order expansion of the model function η. Their method applies to general continuously differentiable models, but requires that the model function have no "parameter-effect" nonlinearities. As shown in related work by Bates and Watts (1980 and 1981), this requirement is not often met by models under a natural parameterization, nor is it easily achieved by reparameterization for most models.

The composite method is accurate so long as the sum-of-squares region is roughly elliptical, even if the ellipse does not coincide with any explicit approximation we might consider. Further, since the composite method approximates the two halves of the sum-of-squares region (for $\tau \geq \hat{\tau}$ and for $\tau \leq \hat{\tau}$) separately, the procedure can perform well even in the presence of substantial asymmetry, unlike any single elliptical approximation. Finally,

an important consideration in the present study, the composite method can be applied to both continuously and piecewise differentiable models. Extensions of the method to accommodate joint confidence regions or models with more than one nonlinear parameter are also possible; the practicality of such procedures remains to be investigated.

estimate by the appropriate t-value. Further, the latter method (F) does not require the derivative of the model function η with respect to the nonlinear parameter τ. However, the linear, t-based confidence intervals (F) and (G) cannot reflect any asymmetry in the likelihood (S) region. As can be seen in Figure 2 of the previous section, this asymmetry is captured by the extreme-tau method (E), and is appropriately incorporated in the composite (C).

For the energy model, the composite approximation (C) was seen in the previous section to be quite accurate, and superior to the Gaussian approximation (G). The difficulty with the Gaussian approximation in general is that any nonlinearity in the model function will make the Gaussian ellipse different from the sum-of-squares contour, even though many types of nonlinearity will leave the sum-of-squares contour itself essentially elliptical.

Hamilton, et al (1981) presented, as an improvement over the Gaussian, an explicit elliptical approximation to the sum-of-squares regions, based on a second-order expansion of the model function η. Their method applies to general continuously differentiable models, but requires that the model function have no "parameter-effect" nonlinearities. As shown in related work by Bates and Watts (1980 and 1981), this requirement is not often met by models under a natural parameterization, nor is it easily achieved by reparameterization for most models.

The composite method is accurate so long as the sum-of-squares region is roughly elliptical, even if the ellipse does not coincide with any explicit approximation we might consider. Further, since the composite method approximates the two halves of the sum-of-squares region (for $\tau > \hat{\tau}$ and for $\tau < \hat{\tau}$) separately, the procedure can perform well even in the presence of substantial asymmetry, unlike any single elliptical approximation. Finally,

an important consideration in the present study, the composite method can be applied to both continuously and piecewise differentiable models. Extensions of the method to accommodate joint confidence regions or models with more than one nonlinear parameter are also possible; the practicality of such procedures remains to be investigated.

APPENDIX

Following is the procedure used to fit Equation (2.2). For more details, see Stram, et al (1982). The terms T_{max} and T_{min} respectively denote the highest and lowest temperatures T_{mj} in the data set.

1) Choose a starting value $\tau^{(0)}$, a maximum number IMAX of iterations i, and checking ranges N1 and N2. (Recommended values are $\tau^{(0)} = 65^{o}F$, IMAX = 10, N1 = 4, N2 = 10.) Set i = 0.

Iterative search by Newton's method

2) Find $\tau^{(i+1)}$ from Equation (3.4).

3) Find $\hat{\tau}$ by one of the following, and choose the corresponding checking range NCHECK:

   a) If $i > IMAX$, set $\hat{\tau} = \tau^{(i)}$ (NCHECK= N2).

   b) If $\tau^{(i+1)} > T_{max}$ or $\tau^{(i+1)} < T_{min}$, set $\hat{\tau} = \tau^{(i)}$ (NCHECK = N2).

   c) If $\tau^{(i+1)} = \tau^{(i-1)}$, set $\hat{\tau}$ = integer between $\tau^{(i)}$ and $\tau^{(i+1)}$ (NCHECK = N1).

   d) If $\tau^{(i+1)} = \tau^{(i)}$, set $\hat{\tau} = \tau^{(i)}$ (NCHECK = N1).

   e) Otherwise, replace i by i + 1 and return to step 2.

Local grid search

4) Set: $\hat{k}$ = integer part of the current estimate $\hat{\tau}$;

   $k_{min} = \max(T_{min}, \hat{k}-NCHECK)$;

   $k_{max} = \min(T_{max}, \hat{k}+NCHECK)$.

5) a) Repeat for $k = k_{min}$ to k_{max}; then go to b):

   If $RSS(k) < RSS(\hat{\tau})$, replace $\hat{\tau}$ by k and replace $\hat{k}$ by k.

   b) If $\hat{\tau} = k_{min}$ or $\hat{\tau} = k_{max}$ continue to c).

   Otherwise go to d).

c) (For the rare cases when 3a) or 3b) gives an initial value $\hat{\tau}$ close to T_{min} or T_{max}.)

For $\hat{k} = \hat{\tau} = k_{min}$ (k_{max}), reset k_{min} (k_{max}) as in Step 4. If the "reset" makes no change, go to d). Otherwise, return to a) for $k = k_{min}$ to $\hat{k}$ ($\hat{k}$ to k_{max}).

d) For $k = \hat{k}-1$ and $k = \hat{k}$:

Find $\hat{\tau}_k$ as $\tau^{(i+1)}$ from Equation (3.4), evaluated for $\tau^{(i)} = k$.

If $\hat{\tau}_k \in [k,k+1]$ and $RSS(\hat{\tau}_k) < RSS(\hat{\tau})$, replace $\hat{\tau}$ by $\hat{\tau}_k$.

6) Computation of remaining parameters

Calculate

$$\hat{\beta} = (SHY/SHH)\Big|_{\hat{\tau}}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{H}(\hat{\tau})$$

$$\hat{\Gamma} = \hat{\alpha} + \hat{\beta}H_0(\hat{\tau}) \quad .$$

A procedure described by Hinkley (1971) for fitting a very similar change point model is equivalent to starting at step (5), and letting k range from $T_{min} + 1$ to $T_{max} - 1$. The full procedure given here greatly reduces the total number of function evaluations required, at the risk of settling on a local minimizer which is not the global minimizer of $RSS(\tau)$. The values N1 and N2 have been chosen to reduce this risk to a very low level (see Goldberg, 1982).

REFERENCES

Bates, D. M. and Watts, D. G. (1980). "Relative Curvature Measures of Nonlinearity." J.R.S.S. B 42 :1-25.

_____ (1981). "Parameter Transformations for Improved Approximate Confidence Regions in Nonlinear Least Squares." Annals of Statistics 9 : 1152-1167.

Dutt, G. S., Lavine, M., Levi, B., and Socolow, R. (1982). "The Modular Retrofit Experiment: Exploring the House Doctor Concept." Princeton University Center for Environmental Studies, Princeton, NJ, Rept. No. 130.

Fels, M. F. and Goldberg, M. L. (1982). "Measuring Household Fuel Consumption on the Standard Living Cycle." Energy 7 : 489-504.

Fels, M. F., Goldberg, M. L., Lavine, M. L., Socolow, R. H., and Abrams, P. S. (1982). "Exploratory Analysis of Oil-Heated Houses." Princeton University Center for Energy and Environmental Studies, Princeton, NJ, Report No. 139.

Goldberg, M. L. (1982). "A Geometrical Approach to Nondifferentiable Regression Models as Related to Methods for Assessing Residential Energy Conservation." Princeton University Center for Energy and Environmental Studies, Princeton, NJ, Rept. No. 142; Ph.D. dissertation for the Department of Statistics.

_____ (1983). "Measures of Nonlinearity for Segmented Regression Models." University of Wisconsin-Madison, Mathematics Research Center, Technical Summary Report #2549.

Hamilton, D. C., Watts, D. G., and Bates, D. M. (1982). "Accounting for Intrinsic Nonlinearity in Nonlinear Regression Parameter Inference Regions." Annals of Statistics 10 : 386-393.

National Oceanic and Atmospheric Administration, U. S. Department of Commerce (monthly publication). "Local Climatological Data, Monthly Summary, Newark, New Jersey."

Stram, D., Dutt, G. S., Fels, M. F., and Goldberg, M. L. (1982). "The Score-keeping Model for Residential Energy Consumption: Procedures and Problems." Princeton University Center for Energy and Environmental Studies, Princeton, NJ, Rept. No. 147.

MLG/PB/Mff/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #2601	2. GOVT ACCESSION NO. AD-A136403	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SIMPLIFIED CONFIDENCE INTERVALS FOR AN AWKWARD NONLINEAR MODEL		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Miriam L. Goldberg, Peter Bloomfield and Margaret F. Fels		8. CONTRACT OR GRANT NUMBER(s) MCS-7927062, Mod. 2 DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics & Probability
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below		12. REPORT DATE November 1983 13. NUMBER OF PAGES 30
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

and

National Science Foundation
Washington, DC 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear, Regression, Confidence Intervals, Energy

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A simple method is developed for approximating confidence intervals for a linear change-point model representing residential energy consumption. The method can be applied to any nonlinear model which is conditionally linear in all but one parameter. The approximation is shown to be more accurate than the Gaussian approximation, and easier to compute than confidence intervals defined by likelihood (sum-of-squares) regions. Computations are described in some detail for the motivating model.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)


DA
FIL